The role of symmetry in neural networks and their Laplacian spectra

de Lange, Siemon C; van den Heuvel, Martijn P; de Reus, Marcel A

published in
NeuroImage
2016

DOI (link to publisher)
10.1016/j.neuroimage.2016.07.051

document version
Peer reviewed version

Title

The role of symmetry in neural networks and their Laplacian spectra

Authors

Siemon C. de Lange, Martijn P. van den Heuvel[*], Marcel A. de Reus[*]

[*] These authors contributed equally to this work.

Affiliation

Brain Center Rudolf Magnus, Department of Psychiatry, University Medical Center Utrecht, The Netherlands

Corresponding Author

Siemon C. de Lange, UMC Utrecht, Department of Psychiatry, A00.241, P.O. Box 85500, 3508 GA Utrecht, The Netherlands

E-mail: s.c.delange@umcutrecht.nl

Abstract

Human and animal nervous systems constitute complexly wired networks that form the infrastructure for neural processing and integration of information. The organization of these neural networks can be analyzed using the so-called Laplacian spectrum, providing a mathematical tool to produce systems-level network fingerprints. In this article, we examine a characteristic central peak in the spectrum of neural networks, including anatomical brain network maps of the mouse, cat and macaque, as well as anatomical and functional network maps of human brain connectivity. We link the occurrence of this central peak to the level of symmetry in neural networks, an intriguing aspect of network organization resulting from network elements that exhibit similar wiring patterns. Specifically, we propose a measure to capture the global level of symmetry of a network and show that, for both empirical networks and network models, the height of the main peak in the Laplacian spectrum is strongly related to node symmetry in the underlying network. Moreover, examination of spectra of duplication-based model networks shows that neural spectra are best approximated using a trade-off between duplication and diversification. Taken together, our results facilitate a better understanding of neural network spectra and the importance of symmetry in neural networks.

Keywords: connectome, neural networks, symmetry, Laplacian, eigenvalue spectrum, duplication

Introduction

Nervous systems constitute complexly wired neural networks. On the macroscale systems level, anatomically segregated brain regions and their interconnecting

pathways together form a complex system of neural nodes and connections, referred to as the macroscale connectome (Sporns et al., 2005). The architecture of connectomes can be examined using tools from graph theory, a branch of mathematics revolving around the formal study of networks (Bullmore and Sporns, 2009; van den Heuvel and Hulshoff Pol, 2010; van den Heuvel et al., 2016). Different from many more typical graph metrics which describe neural systems on the basis of (average) node characteristics (see Rubinov and Sporns (2010) for an overview), the so-called Laplacian spectrum provides a systems-level fingerprint of brain networks and can therefore potentially provide new insights into brain network architecture (Banerjee and Jost, 2007; de Lange et al., 2014; Varshney et al., 2011; Vukadinović et al., 2002). Notably, previous examinations of spectra of neural networks have suggested that neural networks may belong to a special spectral class characterized by small peaks at the start of the spectrum and a pronounced peak in the middle (Banerjee and Jost, 2007; de Lange et al., 2014).

Although some features of Laplacian spectra have been related to conventional network measures such as synchronization (Atay et al., 2006), local clustering (Bauer and Jost, 2012) and modularity (Cheng and Shen, 2010; Shen and Cheng, 2010; Shi and Malik, 2000), several other characteristics remain unexplained. For one, the large central peak in the spectra of neural networks –one of their most characteristic features– is not yet fully understood. Based on theoretical results showing that nodes with identical connectivity patterns increase the elevation precisely in the middle of the spectrum, it has been hypothesized that the central peak may describe nodal symmetry (Banerjee and Jost, 2008), a property noted to be closely related to other features deemed relevant for neural systems such as parallel processing and functional

specialization (Ballard, 1986; Felleman and Van Essen, 1991; Johansen-Berg et al., 2004; Passingham et al., 2002).

In this study, we examine to what extent symmetry shapes neural spectra and whether overlap in the wiring pattern of brain regions can explain the large central peak observed in spectra of neural networks. We examine the relationship between network symmetry, measured by the here introduced duplication coefficient, and the central spectral peak across and within macroscale reconstructions of the mouse, cat, macaque and human connectome, as well as across several non-neural empirical networks (e.g., a food web, a word adjacency network and a book co-purchases network). Next, we investigate this relationship also for commonly used network models. Finally, the interplay between regional symmetry and regional diversity in neural networks is investigated by fitting the spectra of a duplication-based network model to the observed neural spectra.

Materials and Methods

In the next paragraphs, we first describe the adopted network formalism and the examined neural and non-neural networks, followed by a description of the proposed duplication coefficient, performed spectral analysis and used network models.

Network formalism

Laplacian spectra and network symmetry were investigated for network maps of the human, macaque, cat and mouse brain. Networks were mathematically represented by a binary connectivity matrix, formally known as the network's adjacency matrix. The rows and columns of this matrix represented the nodes of the network and an entry in

row *i* and column *j* represented a connection (also referred to as 'edge') between node *i* and node *j*. In the examined neural networks, nodes represented brain regions and connections represented large-scale anatomical pathways between brain regions –or, in case of functional connectivity in the human brain, correlations between spontaneous fluctuations in regional blood oxygen levels. Connectivity matrices were subsequently subjected to graph analysis to assess the Laplacian spectrum and the level of network symmetry (see below).

Neural networks

Human, standard resolution. The human brain connectivity network was obtained from high-quality diffusion-weighted MRI data from 215 subjects, provided in the Q3 data release of the Human Connectome Project (Van Essen et al., 2012; Glasser et al., 2013) as previously described by de Reus and van den Heuvel (2014). Briefly, white matter pathways were reconstructed using generalized q-sampling imaging and streamline tractography (Yeh et al., 2010) and the cortex was parcellated into 68 distinct regions based on FreeSurfer's Desikan-Killiany atlas (Desikan et al., 2006). Two brain regions were considered connected if in at least 60% of the subjects a reconstructed white matter fiber pathway touched both regions (de Reus and van den Heuvel, 2013a). This procedure resulted in a group-averaged connectome map with 68 nodes and a density of 20%.

Human, high resolution. By combining the aforementioned reconstructed white matter pathways with an additional cortical parcellation based on a high-resolution subdivision of the Desikan-Killiany atlas (Cammoun et al., 2012), a second group-averaged connectome map was formed (Hagmann et al., 2008; van den Heuvel and

Sporns, 2011; de Reus and van den Heuvel, 2014). This high-resolution human brain network, again including connections if present in at least 60% of the subjects, comprised 219 nodes and had a density of 5.0%.

Human, functional connectivity. In addition to human anatomical brain networks, also a human functional connectivity network was formed (van den Heuvel et al., 2015). The functional connectivity network described the same 219 regions as the high-resolution anatomical network and was obtained from resting-state functional MRI data as provided in the Q3 data release of the Human Connectome Project. Images were realigned, co-registered with the T1 image, bandpass filtered (0.03 – 0.12 Hz), corrected by means of linear regression for global effects of motion, global signal mean, ventricle signal and white matter signal, and 'scrubbed' to remove potential movement artifacts (Power et al., 2012). Functional connectivity between brain regions was assessed using correlation analysis, computing Pearson's correlation coefficient between time-series composed of the average signal level of all voxels in a cortical region. A group-averaged functional network was formed by averaging the individual functional connectivity networks and keeping connections with an average correlation higher than 0.2, such that the density of the functional network (6.1%) was close to the density of the human high-resolution anatomical network.

Macaque. The macaque connectivity dataset was extracted from the publicly available Collation of Connectivity data for the Macaque (CoCoMac) database, containing information on macroscale white matter pathways as reported in macaque anatomical tracer studies (Stephan et al., 2001). The here used connectivity matrix (previously adopted in the study of Scholtens et al. (2014)) described intrahemispheric

connectivity between 78 non-overlapping cerebral brain regions, parcellated according to a scheme proposed by Felleman and Van Essen (1991). An anatomical tract between two brain regions was included in the matrix if the tract was reported as existing in two-thirds of the reports which investigated the tract. The resulting connectivity matrix for the macaque had 78 nodes and a density of 23%.

Cat. The macroscale cat connectivity network was taken from a study by Scannell et al. (1995), presenting a connectivity dataset based on a comprehensive collation of neural tracing studies in the cat brain. The dataset contained information on projections between 65 non-overlapping brain regions, covering a single hemisphere of the cat cortex (de Reus and van den Heuvel, 2013b). Here, information about the presence of connections was used to form a network with 65 nodes and a density of 35%.

Mouse. The mouse connectivity matrix was based on the recently presented Allen Mouse Brain Connectivity Atlas (http://connectivity.brain-map.org), which describes the axonal projections between brain regions obtained from anterograde tracer examinations performed by a single research group (van den Heuvel and de Reus, 2014; Oh et al., 2014). The connectivity matrix presented by Oh et al. (2014) contained information on the neural wiring of 213 right-hemispheric brain regions, including data about the confidence of every connection. For this study, a connection between two regions was included if the reported α-level of the connection was lower than 0.05. The resulting connectivity matrix had 213 nodes and its density was 12%.

Non-neural empirical networks

In addition to the neural networks described above, also five non-neural empirical networks were examined, representing a food web, a network of frequently combined words, a social network describing book characters from the novel Les Miserables, a social network of interactions between dolphins and a network of co-purchased books on US politics. Table 1 provides a brief overview of these empirical networks and introduces the abbreviations used to refer to them. More detailed descriptions, along with references to the sources from which the networks were obtained, are presented in the Supplementary Materials.

Duplication coefficient

To measure the level of global symmetry in a network, we propose a measure that we will refer to as the *duplication coefficient* of the network. The duplication coefficient measures to what extent network nodes have a 'copy', or, phrased differently, how much nodes look like their most similar peer. The similarity between nodes *i* and *j* is inferred from the overlap between their connectivity pattern, measured by the matching index (Hilgetag et al., 2002):

$$\frac{|N(i) \setminus j \cap N(j) \setminus i|}{|N(i) \setminus j \cup N(j) \setminus i|},$$

where $N(i)$ denotes the set of nodes connected to node *i* (i.e., the topological neighbors of node *i*), $N(i) \setminus j$ are all neighbors of *i* excluding *j* in case nodes *i* and *j* are connected, and $|M|$ denotes the number of nodes in a set *M*. The matching index takes values between 0 (for disjoint connectivity patterns) and 1 (for identical connectivity patterns). The duplication coefficient of an individual node is defined as its maximum similarity to any other node in the network and the duplication coefficient of the

entire network is defined as the average duplication coefficient over all nodes. The rationale for defining nodal duplication in terms of maximum rather than average similarity originates from the mathematical observation that two nodes with identical connectivity patterns give rise to an eigenvalue $\lambda = 1$ in the normalized Laplacian spectrum (see next paragraph), even if the connectivity patterns of those two nodes are completely different from the connectivity patterns of all other nodes in the graph (Banerjee and Jost, 2009). Similarly, the existence of two nodes with highly similar connectivity patterns –as captured by the proposed duplication coefficient– results in an eigenvalue λ near 1 (see Supplementary Materials for exact formulations and mathematical arguments).

Laplacian spectrum

Network spectra were obtained by means of spectral analysis of the so-called normalized Laplacian matrix, as in a previous examination of the spectral properties of neural networks (de Lange et al., 2014). Formally, the normalized Laplacian spectrum is given by the collection of all eigenvalues of the matrix $L=I–D^{-1}A$, where A is the adjacency matrix of the network and D is a diagonal matrix with the degree of the nodes on the diagonal (Chung, 1997). An important feature of the normalized Laplacian matrix L is that all eigenvalues are normalized between 0 and 2, enabling comparison of spectra for differently sized networks (Chung, 1997). For plotting and investigation of Laplacian spectra as continuous curves $\Gamma(x)$ rather than discrete collections of eigenvalues λ_i, eigenvalue densities were convolved with a Gaussian kernel of standard deviation $\sigma = 0.015$,

$$\Gamma(x) = \sum_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{|x-\lambda_i|^2}{2\sigma^2}\right),$$

and then normalized such that the total area under the curve was 1 (Banerjee and Jost, 2007). The height of the main peak in the Laplacian spectrum was taken to be the maximum value of this normalized continuous function.

Relationship between peak height and symmetry

The relationship between the height of the main peak in the Laplacian spectrum and the global level of network symmetry –the existence of such a relationship being the main hypothesis of our study– was examined across and within networks.

Across neural and non-neural empirical networks, the correlation between main peak height and duplication coefficient was considered. Furthermore, the absolute values of the duplication coefficient and central peak height were compared. In addition, within neural networks, the effect of symmetry on the Laplacian spectrum was investigated by comparing how the main peak in the spectrum was altered under two edge rewiring strategies. In the *random* strategy, a small set of nodes (5%, 10%, 15%, 20%) was randomly selected and the edges of these nodes were rewired to randomly chosen new neighbors. The complementing *targeted* strategy was optimized to decrease network symmetry and selected the network with the lowest duplication coefficient from 100 rewired networks produced using the random strategy. For both the random and targeted strategy, 1,000 altered networks were constructed and the rewiring effect on the main peak of the spectrum was assessed by comparing the difference in peak height decrease between random and targeted rewiring.

Network models

To examine how certain network organizations relate to symmetry and main spectral peak height, four different network models were investigated. For each model, 1,000

instances with 219 nodes were constructed using parameters such that the instances matched the human high-resolution connectome (219 nodes, density 5.0%). The main spectral peak height and duplication coefficient distributions of the generated model instances were computed, further examined and compared to the values obtained for the high-resolution human brain network. The investigated network models included the *Erdös-Rényi* (*ER*), *Watts-Strogatz* (*WS*), *Barabási-Albert* (*BA*) and *geometric* model:

Erdös-Rényi (ER) model. ER random networks are completely random networks with no other specified structure than their size and density (Erdös and Rényi, 1959) and were obtained by randomly placing 1095 connections between 219 nodes, resulting in an average node degree of 10 (i.e., nodes on average have 10 connections). ER random networks show no community structure, low levels of clustering and short path lengths.

Watts-Strogatz (WS) model. WS model networks were obtained by rewiring 12% of the edges of a lattice ring network of nodes with degree 10 (Watts and Strogatz, 1998). WS model networks display a small-world organization, characterized by high levels of local clustering and short path lengths. A combination of strong clustering and short path lengths is also reported for neural networks (Bassett and Bullmore, 2006; Bullmore and Sporns, 2009; Hagmann et al., 2008; van den Heuvel et al., 2008). The rewiring percentage was chosen such that the clustering coefficient of the generated small-world networks matched that of the human connectome.

Barabási-Albert (BA) model. BA model networks were grown from a seed network with 5 fully connected nodes by iteratively introducing new nodes to the network. Each new node was randomly connected to 5 already present nodes according to a degree-biased probability function (Barabási and Albert, 1999). The probability for a connection of a new node *i* to be attached to an already present node *j* was

$$P_j = \frac{k_j}{\sum_l k_l},$$

where k_j is the degree of node *j* and the sum is over all existing nodes *l*. BA model networks have scale-free degree distributions and thus comprise small numbers of highly connected hub nodes (Barabási and Albert, 1999), a feature also observed across neural networks (Achard et al., 2006; van den Heuvel and Sporns, 2011; Sporns et al., 2007; Zamora-López et al., 2010).

Geometric model. In the geometric model, networks grew in a similar manner as in the BA model, but nodes were randomly embedded in a three-dimensional cube and newly added nodes were connected to the closest nodes in Euclidean space. This cost-minimizing model helped to understand how symmetry and the central spectral peak interact with spatial optimal wiring, a principle suggested to be involved in establishing the connectivity pattern of neural networks (Bassett et al., 2010; Bullmore and Sporns, 2012; Collin et al., 2014; van den Heuvel et al., 2012; Sporns et al., 2004; Vértes et al., 2012).

Duplication model

A duplication-based model in which symmetry naturally emerges from iterative (partial) node duplication (Bhan et al., 2002) was implemented to investigate the extent to which the architecture and spectra of neural networks may be driven by

symmetry patterns. Model networks grew from a seed network by duplication of randomly selected nodes. Duplication of a node, referred to as the parent node, involved the introduction of a new node to the network that was connected to all neighbors of the randomly chosen parent node. One connection of the new node was rewired such that the new node was connected to its parent, while the other connections of the new node were rewired with probability 1 - γ to nodes not in the neighborhood of the parent node, diversifying the node's connectivity profile. Seed networks were randomly wired networks with average degree 2 and the size of a seed network was chosen such that the density of the seed network matched the density of the neural network that was to be approximated.

To empirically establish the balance between regional diversity and regional symmetry in neural networks (modeled by parameter γ), a grid search was applied to find the values of γ for which the spectra of the duplication model best resembled the examined neural spectra. Resemblance was measured using the concept of spectral distance (Gu et al., 2016), estimating the difference between two spectral plots Γ_1 and Γ_2 by

$$D(\Gamma_1, \Gamma_2) = \int_0^2 |\Gamma_1(x) - \Gamma_2(x)|\, dx,$$

where the integral was discretely evaluated using intervals of length 0.001. In the grid search, the performance of parameter γ on the interval [0, 1] was evaluated with steps of 0.01 using the spectra of 100 model realizations.

Results

Spectral peak height and symmetry across empirical networks

Duplication coefficient scores showed a clear association with central peak height across empirical networks ($r = 0.86$, $p < 0.001$, Figure 1). The duplication coefficient of the examined neural networks revealed to range between 0.48 and 0.69 and their central peak height between 1.9 and 3.6, causing the neural networks to form a cluster positioned in between the examined non-neural empirical networks. Non-neural empirical networks showed a distinction between, on the one hand, the dolphins, words and books networks, having lower duplication coefficients (between 0.33 and 0.46) and smaller central peaks (between 1.4 and 1.7) than neural networks and, on the other hand, the food web and Les Miserables network, reporting high levels of duplication (0.68 and 0.75 respectively) and a high central peak (5.1 and 4.9). The high spectral peak and symmetry level observed for the food web is in agreement with earlier observations demonstrating overlap between food web and neural network spectra (de Lange et al., 2014), as well as with suggestions of connectivity similarity playing a major role in the organization of food webs (Caldarelli et al., 1998; Drossel and McKane, 2002).

Effects of targeted and non-targeted rewiring

The influence of symmetry on the main spectral peak was further investigated within neural networks by examining targeted and non-targeted edge rewiring. For all neural networks, targeted rewiring of edges (optimized to decrease network symmetry) induced a stronger decrease of the main spectral peak than random rewiring (see Figure 2 for 10% rewiring and Supplementary Figure 1 for 5%, 15% and 20% rewiring; all $p < 0.001$, *t*-tests, Bonferroni corrected). The largest effect was observed for the human anatomical brain network at standard resolution, where targeted rewiring decreased the height of the main central peak 1.5 times more than non-

targeted rewiring (human, high resolution: ×1.1; macaque: ×1.2; mouse, cat and human functional connectivity: ×1.3).

Validation using network models

The existence of a relationship between duplication coefficient and spectral peak height was verified using four network models. In comparison to the high-resolution human brain network to which the model networks were matched, ER random and BA scale-free model networks showed lower peak heights (respectively 42% and 40% lower than the human network, $p < 0.001$) and lower duplication coefficients (66% and 57% lower, $p < 0.001$, Figure 3). Instances of the WS small-world model and geometric model showed similarly high levels of symmetry as seen in the human high-resolution network, but generally lower peak heights (average height for WS 1.9, average height for geometric 1.7) than the human network (height 2.0). The relationship between duplication coefficient and peak height observed within and between the network models was in line with the trend seen for the empirical networks: all network models revealed a significant linear association between peak height and duplication coefficient across the 1,000 generated model instances (ER: $r = 0.13$, WS: $r = 0.64$, BA: $r = 0.30$, geometric: $r = 0.40$, with $p < 0.001$ for all models) and also the relative arrangement of the four populations was found to be in support of such an association (Figure 3).

Modeling neural spectra using nodal duplication

As indicated by the observed duplication coefficient scores for neural networks (ranging between 0.48 and 0.69), perfect symmetry is rare in neural networks. That is, the connectivity overlap between a brain region and its most similar peer is generally

only partial. To provide insight in the balance between symmetry and diversity in the connectivity of brain regions, the spectrum of each neural network was approximated by the spectra of duplication-based model networks using a spectral fitting procedure. The best-fitting model spectra and the associated empirically established optimal parameters γ of the duplication model are shown in Figure 4. Values of γ, reflecting the expected percentage of connections that was copied from the parent during duplication and thus providing a symmetry versus diversity estimate (with 1 reflecting total duplication and 0 reflecting maximum diversity), were found to range from 0.74 for the best-fitting model of the mouse connectome, to 0.93 for the best-fitting model of the cat connectome. These relatively high values for γ indicate that duplication played a prominent role during the growth of the best-fitting model networks, but also that a certain amount of regional diversification (established by rewiring 7-26% of the connections after duplication) was essential to obtain spectra that resemble those of neural networks.

Figure 5 shows the duplication coefficient and central peak height for instances of the duplication-based network model generated using different values of γ. As in the analysis of the four common network models, all instances were matched (in terms of density and network size) to the high-resolution human brain network. In line with our expectations, both central peak height and duplication coefficient increased with γ (Figure 5). Furthermore, even though the duplication coefficient played no role in the spectrum-based fitting procedure, the duplication coefficients (mean/std: 0.54/0.018) of the best-fitting model networks, corresponding to $\gamma = 0.83$, were found to be approximately equal to the duplication coefficient (0.52) of the high-resolution human connectome (showing no significant difference, $p = 0.21$).

Discussion

This study describes the role of symmetry in the architecture of neural networks as revealed through their Laplacian spectra. Investigation of both empirical and model networks demonstrates that the large peak observed in the center of neural network spectra is strongly influenced by the high level of node symmetry in neural networks. Such a relation is in line with earlier theoretical results identifying the occurrence of perfectly symmetric nodes as a source for eigenvalues positioned exactly in the middle of the spectral domain (Banerjee and Jost, 2008; Yadav and Jalan, 2015).

Three lines of evidence support the notion of node symmetry in neural networks to be an important factor in shaping the spectral peak. First, across both neural and non-neural networks, the height of the central peak correlated to the networks' duplication coefficient (Figure 1). Second, this relationship between duplication coefficient and spectral peak height was also observed across and within commonly used network models (Figure 3), with significant correlations between duplication coefficient and spectral peak height being observed among realizations of each model. Third, perturbations (in the form of edge rewiring) targeting network symmetry revealed to have a significantly larger impact on the spectral peak height of neural networks than random perturbations (Figure 2).

Considering possible mechanisms behind the occurrence of symmetry in neural networks, we note that geometric and Watts-Strogatz small-world model networks show high duplication coefficients compared to both neural networks and other network models. In geometric model networks, proximate nodes are likely to have

similar wiring patterns, increasing the duplication coefficient (Song et al., 2014). Interestingly, previous studies suggested that brain networks are shaped by a trade-off between minimization of wiring length or 'cost' –being the driving mechanism behind the geometric model– and establishing topological beneficial factors such as high local clustering, both of which may contribute to the observed node symmetry in brain networks (Bassett et al., 2010; Betzel et al., 2015; Bullmore and Sporns, 2012; van den Heuvel et al., 2016; Kaiser and Hilgetag, 2006; Raj and Chen, 2011; Vértes et al., 2012).

The elevated symmetry in WS model networks is attributable to their high clustering coefficient (directly induced by the model's design), which is a normalized measure of the number of triangles in the network, referring to subgroups of three mutually connected nodes. Since any two nodes from such a triangle have at least one projection in common, the presence of triangles induces a certain level of node symmetry. However, the concepts of clustering and duplication are inherently different, as is illustrated by the food web, which showed the highest spectral peak among all empirical networks and a high duplication coefficient (0.68 compared to 0.52 in the high-resolution human connectome), but a low clustering coefficient (0.29 compared to 0.48 in the high-resolution connectome). This combination of low clustering and high symmetry is a consequence of the nature of the food web, in which species that feed on the same prey (or are eaten by the same predators) are likely to be at the same level in the food chain and hence unlikely to have a predator-prey relation with each other. In such a situation where similar nodes are often unconnected, high overlap between the wiring pattern of nodes will generate many 'open' triangles and thus decrease the clustering coefficient of the network.

Underscoring this conceptual difference between clustering and duplication, post-hoc analysis indeed showed that the clustering coefficient and peak height of the investigated empirical networks were not correlated ($p = 0.51$). Furthermore, when regressing out the effects of clustering on central peak height, the central peak height residue and duplication coefficient remained significantly associated ($r = 0.73$, $p = 0.011$, see Supplementary Figure 2).

Taking a closer look at the duplication coefficient values observed for the neural networks, the occurrence of brain areas with completely identical wiring appears to be rare. Indeed, a post-hoc analysis confirmed that none of the neural networks possess perfectly symmetric brain regions and that only few brain regions have near perfect 'copies' (indicated by a node-wise duplication coefficient greater than 0.9; cat: 3.1% of the regions; human, standard resolution: 5.9%; human, functional connectivity: 2.7%; other networks: 0%). That is, most brain regions exhibit only partially overlapping wiring patterns, suggesting a balance between connectivity symmetry and diversity in neural systems. With brain areas having anatomical connectivity overlap being likely to show functional similarity (Passingham et al., 2002), balancing between symmetry and diversity is potentially linked to functional specialization. In this light, symmetry and diversity may have complementary beneficial effects. Connectivity diversification may support effective global communication and information integration in the network (van den Heuvel and Sporns, 2011; Markov et al., 2013). On the other hand, connectivity symmetry can provide redundancy and robustness (MacArthur et al., 2008) and may facilitate parallel processing (Ballard, 1986). Fitting duplication-based model networks to the neural networks showed that the architecture of neural networks is best approximated if the fraction of connections

copied during duplication ranges between 0.74 and 0.93, confirming the relevance of both symmetry and diversity in neural systems.

Although we here focus on the height of the main peak in neural spectra, the location of the main spectral peak is likely to also reflect important aspects of network topology. As previously discussed (de Lange et al., 2014), the location of small peaks at the start of the spectrum has been shown to describe a network's modularity structure (Shi and Malik, 2000), the location of peaks on the right of the central peak relates to the so-called bipartiteness of a network (Bauer and Jost, 2012), and the occurrence of network motifs may generate peaks at specific locations in the spectrum (Banerjee and Jost, 2008). Interestingly, also the location of the here studied central peak –typically expected in the middle of the spectral domain– is not fixed and in fact showed to vary across species (see Figure 4). The central peak location in the anatomical human network spectra (both at normal and high resolution) deviated most from 1, showing a clear shift to the right. The precise origin of this shift is speculation at this time, but prior theoretical results (replicated in the Supplementary Materials) showed that the eigenvalue associated with two symmetric, but mutually connected (and hence not perfectly symmetric), nodes is equal to $\lambda = 1 + 1/d$, where d is the degree of the nodes, providing a possible mechanism for the shift. In line with this suggestion, a post-hoc analysis showed a significant correlation between the percentage of nodes connected to its most similar peer and peak height location across neural networks ($r = 0.87$, $p = 0.025$, see Supplementary Figure 3).

When interpreting the findings of our study, it should be noted that the examined neural network maps resulted from a variety of reconstruction methodologies, limiting

the possibility to compare the relative positions of the different species in the duplication coefficient versus peak height plot (Figure 1) (van den Heuvel et al., 2016). Indeed, reconstruction methodology is likely to have an effect on elementary properties such as network density, which in turn have impact on the shape of the spectrum (Chung et al., 2003). In this context, it should also be noted that some of the examined brain networks cover a single hemisphere (e.g., the cat, macaque and mouse), while others comprised both hemispheres (e.g., variants of the human brain network). However, a post-hoc analysis on the bihemispheric networks showed that the most similar peer of a node is often situated in the same hemisphere (82% for the standard-resolution anatomical human connectome, 91% for the high-resolution version and 72% for functional connectivity), suggesting that whether a neural network covers one or two hemispheres may not be a key determinant for the observed duplication coefficient. Indeed, the average duplication coefficient and central spectral peak height of the right and left hemisphere considered separately was close to the original values (standard-resolution: duplication coefficient 0.64 (original 0.62), central peak height 2.7 (2.6); high-resolution: 0.52 (0.52), 1.9 (1.9); functional connectivity: 0.54 (0.53), 2.0 (2.4)).

Even though our findings demonstrate an association between the introduced duplication coefficient and spectral peak height, it is important to stress that correlation analysis cannot substantiate the existence of a causal or direct relationship. Interestingly, considering Figure 3 and Figure 5, the relationship between duplication coefficient and central peak height might actually be more complex than the now explored linear relation, with a quadratic relationship providing a significantly better fit for the values in Figures 3 and 5 (both $p < 0.001$, F-tests). Figure 3 and Figure 5

further suggest that the strength of the association between central peak height and duplication becomes stronger for increasing levels of symmetry. Indeed, the correlation coefficient as determined within network groups corresponding to a particular network model (Figure 3) or value of γ (Figure 5) was higher for groups with a higher average duplication coefficient ($p < 0.001$, *t*-tests using bootstrapping, 1,000 samples, Bonferroni corrected), suggesting other sources of variation are more apparent when the duplication coefficient is low.

One potential source of variation that is not incorporated in the duplication coefficient is given by the existence of symmetries more complex than overlap in node wiring (MacArthur and Anderson, 2006). Indeed, the notion of perfect node symmetry –in which the label of two nodes can be permuted without altering the network structure– can be extended to higher-level symmetries, including the scenario in which two groups of nodes have the same combined wiring pattern (Yadav and Jalan, 2015). As for node duplication, such higher-level symmetries have been shown to induce a peak in the middle of the spectrum (Dorogovtsev et al., 2003; Vukadinović et al., 2002; Yadav and Jalan, 2015), suggesting that also (approximate) higher-level symmetries can contribute to the central spectral peak. Higher-level network symmetries have also been linked to aspects of network dynamics, such as controllability and synchronization (O'Clery et al., 2013; Pecora et al., 2014; Sorrentino et al., 2016). Future attempts to quantify the presence of (approximate) higher-level symmetries should reveal the relevance of such symmetries in neural networks and their spectra and might further contribute to our understanding of neural spectra by relating network dynamics to the size and shape of the central spectral peak.

Another future direction would be the adoption of the Laplacian spectrum as an instrument to describe network abnormalities in brain disorders. Relating the characteristic central peak in the Laplacian spectrum to node symmetry, this study shows how particular deviations in spectral fingerprints may correspond to concrete aspects of network architecture relevant for the study of brain disorders. For instance, a recent study suggested symmetry between brain regions, in combination with other measures, as a classifier for distinguishing between brain connectivity of normal healthy controls, people with Alzheimer's disease and people with mild cognitive impairment (Prasad et al., 2015), which is –in view of our current findings– also likely to be reflected in the associated Laplacian spectra. Although the height of the central peak in network spectra can now be captured by the duplication coefficient, disease-related deviations in the central peak and other parts of the spectrum may in the future reveal systems-level changes, including higher-level symmetry disruptions, that can currently not be described or detected by standard network metrics.

Acknowledgements

Data were provided in part by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University. MPvdH was supported by a VIDI grant of the Dutch Research Council (NWO VIDI-452-16-015) and a Fellowship of MQ.

References

Achard, S., Salvador, R., Whitcher, B., Suckling, J., and Bullmore, E. (2006). A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. *J. Neurosci.* 26, 63–72. doi:10.1523/jneurosci.3874-05.2006.

Atay, F., Biyikoglu, T., and Jost, J. (2006). Network synchronization: Spectral versus statistical properties. *Phys. D Nonlinear Phenom.* 224, 35–41. doi:10.1016/j.physd.2006.09.018.

Ballard, D. H. (1986). Cortical connections and parallel processing: Structure and function. *Behav. Brain Sci.* 9, 67–90. doi:10.1017/S0140525X00021555.

Banerjee, A., and Jost, J. (2007). Spectral plots and the representation and interpretation of biological data. *Theory Biosci.* 126, 15–21. doi:10.1007/s12064-007-0005-9.

Banerjee, A., and Jost, J. (2008). On the spectrum of the normalized graph Laplacian. *Linear Algebra Appl.* 428, 3015–3022. doi:10.1016/j.laa.2008.01.029.

Banerjee, A., and Jost, J. (2009). Graph spectra as a systematic tool in computational biology. *Discret. Appl. Math.* 157, 2425–2431. doi:10.1016/j.dam.2008.06.033.

Barabási, A. L., and Albert, R. (1999). Emergence of Scaling in Random Networks. *Science.* 286, 509–512. doi:10.1126/science.286.5439.509.

Bassett, D. S., and Bullmore, E. (2006). Small-World Brain Networks. *Neurosci.* 12, 512–523. doi:10.1177/1073858406293182.

Bassett, D. S., Greenfield, D. L., Meyer-Lindenberg, A., Weinberger, D. R., Moore, S. W., and Bullmore, E. T. (2010). Efficient physical embedding of topologically complex information processing networks in brains and computer circuits. *PLoS Comput. Biol.* 6, e1000748. doi:10.1371/journal.pcbi.1000748.

Bauer, F., and Jost, J. (2012). Bipartite and neighborhood graphs and the spectrum of the normalized graph Laplacian. *Commun. Anal. Geom.* 21, 787–845.

doi:10.4310/cag.2013.v21.n4.a2.

Betzel, R. F., Avena-Koenigsberger, A., Goñi, J., He, Y., de Reus, M. A., Griffa, A., et al. (2015). Generative models of the human connectome. *Neuroimage*. doi:10.1016/j.neuroimage.2015.09.041.

Bhan, A., Galas, D. J., and Dewey, T. G. (2002). A duplication growth model of gene expression networks. *Bioinformatics* 18, 1486–1493. doi:10.1093/bioinformatics/18.11.1486.

Bullmore, E., and Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* 10, 312. doi:10.1038/nrn2618.

Bullmore, E., and Sporns, O. (2012). The economy of brain network organization. *Nat. Rev. Neurosci.* 13, 336–49. doi:10.1038/nrn3214.

Caldarelli, G., Higgs, P., and McKane, A. (1998). Modelling Coevolution in Multispecies Communities. *J. Theor. Biol.* 193, 345–358. doi:10.1006/jtbi.1998.0706.

Cammoun, L., Gigandet, X., Meskaldji, D., Thiran, J. P., Sporns, O., Do, K. Q., et al. (2012). Mapping the human connectome at multiple scales with diffusion spectrum MRI. *J. Neurosci. Methods* 203, 386–97. doi:10.1016/j.jneumeth.2011.09.031.

Cheng, X. Q., and Shen, H. W. (2010). Uncovering the community structure associated with the diffusion dynamics of networks. *J. Stat. Mech. Theory Exp.* 2010, P04024. doi:10.1088/1742-5468/2010/04/p04024.

Chung, F. R. K. (1997). *Spectral Graph Theory*. Fresno, CA: American Mathematical Society.

Chung, F. R. K., Lu, L., and Vu, V. (2003). Spectra of random graphs with given expected degrees. *Proc. Natl. Acad. Sci.* 100, 6313–8. doi:10.1073/pnas.0937490100.

Collin, G., Sporns, O., Mandl, R. C. W., and van den Heuvel, M. P. (2014). Structural and functional aspects relating to cost and benefit of rich club organization in the human cerebral cortex. *Cereb. Cortex* 24, 2258–67. doi:10.1093/cercor/bht064.

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., et al. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* 31, 968–80. doi:10.1016/j.neuroimage.2006.01.021.

Dorogovtsev, S. N., Goltsev, A. V., Mendes, J. F. F., and Samukhin, A. N. (2003). Spectra of complex networks. *Phys. Rev. E* 68, 046109. doi:10.1103/physreve.68.046109.

Drossel, B., and McKane, A. J. (2002). "Modelling food webs" in *Handbook of Graphs and Networks*, eds. S. Bornholdt and H. G. Schuster (Weinheim, FRG: Wiley-VCH Verlag GmbH & Co. KGaA), 218–247. doi:10.1002/3527602755.

Erdös, P., and Rényi, A. (1959). On random graphs. *Publ. Math. Debrecen* 6, 290–297.

Van Essen, D. C., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T. E. J., Bucholz, R., et al. (2012). The Human Connectome Project: a data acquisition perspective. *Neuroimage* 62, 2222–31. doi:10.1016/j.neuroimage.2012.02.018.

Felleman, D. J., and Van Essen, D. C. (1991). Distributed Hierarchical Processing in the Primate Cerebral Cortex. *Cereb. Cortex* 1, 1–47. doi:10.1093/cercor/1.1.1.

Glasser, M. F., Sotiropoulos, S. N., Wilson, J. A., Coalson, T. S., Fischl, B., Andersson, J. L., et al. (2013). The minimal preprocessing pipelines for the Human Connectome Project. *Neuroimage* 80, 105–24. doi:10.1016/j.neuroimage.2013.04.127.

Gu, J., Jost, J., Liu, S., and Stadler, P. F. (2016). Spectral classes of regular, random, and empirical graphs. *Linear Algebra Appl.* 489, 30–49. doi:10.1016/j.laa.2015.08.038.

Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C. J., Wedeen, V. J., et al. (2008). Mapping the Structural Core of Human Cerebral Cortex. *PLoS Biol* 6, e159. doi:10.1371/journal.pbio.0060159.

van den Heuvel, M. P., Bullmore, E. T., and Sporns, O. (2016). Comparative Connectomics. *Trends Cogn. Sci.* doi:10.1016/j.tics.2016.03.001.

van den Heuvel, M. P., and Hulshoff Pol, H. E. (2010). Exploring the brain network: A review on resting-state fMRI functional connectivity. *Eur. Neuropsychopharmacol.* 20, 519–534. doi:10.1016/j.euroneuro.2010.03.008.

van den Heuvel, M. P., Kahn, R. S., Goñi, J., and Sporns, O. (2012). High-cost, high-capacity backbone for global brain communication. *Proc. Natl. Acad. Sci. U. S. A.* 109, 11372–7. doi:10.1073/pnas.1203593109.

van den Heuvel, M. P., and de Reus, M. A. (2014). Chasing the dreams of early connectionists. *ACS Chem. Neurosci.* 5, 491–3. doi:10.1021/cn5000937.

van den Heuvel, M. P., Scholtens, L. H., Feldman Barrett, L., Hilgetag, C. C., and de Reus, M. A. (2015). Bridging Cytoarchitectonics and Connectomics in Human Cerebral Cortex. *J. Neurosci.* 35, 13943–13948. doi:10.1523/jneurosci.2630-15.2015.

van den Heuvel, M. P., and Sporns, O. (2011). Rich-Club Organization of the Human Connectome. *J. Neurosci.* 31, 15775–15786. doi:10.1523/jneurosci.3539-11.2011.

van den Heuvel, M. P., Stam, C. J., Boersma, M., and Hulshoff Pol, H. E. (2008). Small-world and scale-free organization of voxel-based resting-state functional connectivity in the human brain. *Neuroimage* 43, 528–539. doi:10.1016/j.neuroimage.2008.08.010.

Hilgetag, C. C., Kötter, R., Stephan, K. E., and Sporns, O. (2002). Computational Methods for the Analysis of Brain Connectivity. 295–335. doi:10.1007/978-1-59259-275-3_14.

Johansen-Berg, H., Behrens, T. E. J., Robson, M. D., Drobnjak, I., Rushworth, M. F. S., Brady, J. M., et al. (2004). Changes in connectivity profiles define functionally distinct regions in human medial frontal cortex. *Proc. Natl. Acad. Sci. U. S. A.* 101, 13335–40. doi:10.1073/pnas.0403743101.

Kaiser, M., and Hilgetag, C. C. (2006). Nonoptimal component placement, but short processing paths, due to long-distance projections in neural systems. *PLoS Comput. Biol.* 2, e95. doi:10.1371/journal.pcbi.0020095.

de Lange, S. C., de Reus, M. A., and van den Heuvel, M. P. (2014). The Laplacian spectrum of neural networks. *Front. Comput. Neurosci.* 7, 189.

doi:10.3389/fncom.2013.00189.

MacArthur, B. D., and Anderson, J. W. (2006). Symmetry and Self-Organization in Complex Systems. Available at: http://arxiv.org/abs/cond-mat/0609274.

MacArthur, B. D., Sánchez-García, R. J., and Anderson, J. W. (2008). Symmetry in complex networks. *Discret. Appl. Math.* 156, 3525–3531. doi:10.1016/j.dam.2008.04.008.

Markov, N. T., Ercsey-Ravasz, M., Lamy, C., Ribeiro Gomes, A. R., Magrou, L., Misery, P., et al. (2013). The role of long-range connections on the specificity of the macaque interareal cortical network. *Proc. Natl. Acad. Sci. U. S. A.* 110, 5187–92. doi:10.1073/pnas.1218972110.

O'Clery, N., Yuan, Y., Stan, G. B., and Barahona, M. (2013). Observability and coarse graining of consensus dynamics through the external equitable partition. *Phys. Rev. E* 88. doi:10.1103/PhysRevE.88.042805.

Oh, S. W., Harris, J. A., Ng, L., Winslow, B., Cain, N., Mihalas, S., et al. (2014). A mesoscale connectome of the mouse brain. *Nature* 508, 207–14. doi:10.1038/nature13186.

Passingham, R. E., Stephan, K. E., and Kötter, R. (2002). The anatomical basis of functional localization in the cortex. *Nat. Rev. Neurosci.* 3, 606–16. doi:10.1038/nrn893.

Pecora, L. M., Sorrentino, F., Hagerstrom, A. M., Murphy, T. E., and Roy, R. (2014). Cluster synchronization and isolated desynchronization in complex networks with symmetries. *Nat. Commun.* 5, 4079. doi:10.1038/ncomms5079.

Power, J. D., Barnes, K. a., Snyder, A. Z., Schlaggar, B. L., and Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage* 59, 2142–2154. doi:10.1016/j.neuroimage.2011.10.018.

Prasad, G., Joshi, S. H., Nir, T. M., Toga, A. W., and Thompson, P. M. (2015). Brain connectivity and novel network measures for Alzheimer's disease classification. *Neurobiol. Aging* 36 Suppl 1, S121–31. doi:10.1016/j.neurobiolaging.2014.04.037.

Raj, A., and Chen, Y. (2011). The wiring economy principle: connectivity determines anatomy in the human brain. *PLoS One* 6, e14832. doi:10.1371/journal.pone.0014832.

de Reus, M. A., and van den Heuvel, M. P. (2013a). Estimating false positives and negatives in brain networks. *Neuroimage* 70, 402–9. doi:10.1016/j.neuroimage.2012.12.066.

de Reus, M. A., and van den Heuvel, M. P. (2013b). Rich club organization and intermodule communication in the cat connectome. *J. Neurosci.* 33, 12929–39. doi:10.1523/jneurosci.1448-13.2013.

de Reus, M. A., and van den Heuvel, M. P. (2014). Simulated rich club lesioning in brain networks: a scaffold for communication and integration? *Front. Hum. Neurosci.* 8, 647. doi:10.3389/fnhum.2014.00647.

Rubinov, M., and Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. *Neuroimage* 52, 1059–69. doi:10.1016/j.neuroimage.2009.10.003.

Scannell, J. W., Blakemore, C., and Young, M. P. (1995). Analysis of connectivity in the cat cerebral cortex. *J. Neurosci.* 15, 1463–1483.

Scholtens, L. H., Schmidt, R., de Reus, M. A., and van den Heuvel, M. P. (2014). Linking macroscale graph analytical organization to microscale neuroarchitectonics in the macaque connectome. *J. Neurosci.* 34, 12192–205. doi:10.1523/jneurosci.0752-14.2014.

Shen, H. W., and Cheng, X. Q. (2010). Spectral methods for the detection of network community structure: a comparative analysis. *J. Stat. Mech. Theory Exp.* 2010, P10020. doi:10.1088/1742-5468/2010/10/p10020.

Shi, J., and Malik, J. (2000). Normalized cuts and image segmentation. *Pattern Anal. Mach. Intell. IEEE Trans.* 22, 888–905. doi:10.1109/34.868688.

Song, H. F., Kennedy, H., and Wang, X. J. (2014). Spatial embedding of structural similarity in the cerebral cortex. *Proc. Natl. Acad. Sci.* 111, 16580–16585. doi:10.1073/pnas.1414153111.

Sorrentino, F., Pecora, L. M., Hagerstrom, A. M., Murphy, T. E., and Roy, R. (2016). Complete characterization of the stability of cluster synchronization in complex dynamical networks. *Sci. Adv.* 2, e1501737. doi:10.1126/sciadv.1501737.

Sporns, O., Chialvo, D., Kaiser, M., and Hilgetag, C. (2004). Organization, development and function of complex brain networks. *Trends Cogn. Sci.* 8, 418–425. doi:10.1016/j.tics.2004.07.008.

Sporns, O., Honey, C. J., and Kötter, R. (2007). Identification and classification of hubs in brain networks. *PLoS One* 2, e1049. doi:10.1371/journal.pone.0001049.

Sporns, O., Tononi, G., and Kötter, R. (2005). The human connectome: A structural description of the human brain. *PLoS Comput. Biol.* 1, e42. doi:10.1371/journal.pcbi.0010042.

Stephan, K. E., Kamper, L., Bozkurt, A., Burns, G. A., Young, M. P., and Kötter, R. (2001). Advanced database methodology for the Collation of Connectivity data on the Macaque brain (CoCoMac). *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 356, 1159–1186. doi:10.1098/rstb.2001.0908.

Varshney, L. R., Chen, B. L., Paniagua, E., Hall, D. H., and Chklovskii, D. B. (2011). Structural Properties of the Caenorhabditis elegans Neuronal Network. *PLoS Comput Biol* 7, e1001066. doi:10.1371/journal.pcbi.1001066.

Vértes, P. E., Alexander-Bloch, A. F., Gogtay, N., Giedd, J. N., Rapoport, J. L., and Bullmore, E. T. (2012). Simple models of human brain functional networks. *Proc. Natl. Acad. Sci. U. S. A.* 109, 5868–73. doi:10.1073/pnas.1111738109.

Vukadinović, D., Huang, P., and Erlebach, T. (2002). "On the Spectrum and Structure of Internet Topology Graphs" in *Innovative Internet Computing Systems* Lecture Notes in Computer Science., eds. H. Unger, T. Böhme, and A. Mikler (Springer Berlin Heidelberg), 83–95. doi:10.1007/3-540-48080-3_8.

Watts, D. J., and Strogatz, S. H. (1998). Collective dynamics of "small-world" networks. *Nature* 393, 440–442. doi:10.1038/30918.

Yadav, A., and Jalan, S. (2015). Origin and implications of zero degeneracy in networks spectra. *Chaos* 25, 043110. doi:10.1063/1.4917286.

Yeh, F. C., Wedeen, V. J., and Tseng, W. Y. I. (2010). Generalized q-sampling imaging. *IEEE Trans. Med. Imaging* 29, 1626–35.

doi:10.1109/tmi.2010.2045126.

Zamora-López, G., Zhou, C., and Kurths, J. (2010). Cortical hubs form a module for multisensory integration on top of the hierarchy of cortical networks. *Front. Neuroinform.* 4, 1. doi:10.3389/neuro.11.001.2010.

Figure captions

Figure 1. Duplication coefficient and central peak height in empirical networks. Duplication coefficient (measuring network symmetry) and the height of the central peak in the Laplacian spectrum were correlated across empirical networks, including six neural and five non-neural networks ($r = 0.86$, $p < 0.001$). This correlation suggests that network symmetry is an important factor contributing to the high central peaks which have been previously shown to be a key characteristic of neural network spectra (de Lange et al., 2014).

Figure 2. Effect of edge rewiring on height of the central spectral peak. Targeted rewiring of edges (10%, see Supplementary Figure 1 for other levels of rewiring) optimized to reduce the duplication coefficient of neural networks was found to induce a significantly larger decrease of the central spectral peak than randomly rewiring edges (all $p < 0.001$, *t*-tests, Bonferroni corrected).

Figure 3. Duplication coefficient and central peak height in model networks. The duplication coefficients and central spectral peak heights across four different network models underscored the existence of an association between these measures. Significant associations were also found between the instances of each model separately (inserts, *x*-axis: duplication coefficient, *y*-axis: central peak height; all $p < 0.001$, Bonferroni corrected).

Figure 4. Approximation of neural spectra using duplication-based model networks. The optimal fraction γ of copied connections during the growth of

duplication-based model networks ($\gamma = 1$ being complete duplication and $\gamma = 0$ complete diversification) was empirically determined by spectrally fitting the model networks to the connectivity maps of the **(a)** cat (optimal $\gamma = 0.93$), **(b)** macaque ($\gamma = 0.87$), **(c)** mouse ($\gamma = 0.74$), **(d)** human ($\gamma = 0.89$), **(e)** high-resolution human ($\gamma = 0.83$) and **(f)** functional human ($\gamma = 0.87$) brain networks. The reported optimal values of γ suggest a balance in the interplay between connectivity similarity and diversity of brain regions.

Figure 5. Effect of duplication-based model parameter γ on duplication coefficient and central peak height. Increasing the local symmetry during the growth of the duplication-based model networks, reflected by the parameter γ, increased both the duplication coefficient and central peak height. Comparison with the high-resolution human connectome showed that duplication-based model networks generated with optimal parameter $\gamma = 0.83$ match the human connectome in terms of both central peak height and duplication coefficient.

Supplementary Figure 1. Effect of 5%, 15% and 20% edge rewiring on the height of the central spectral peak. For all edge rewiring percentages and all networks, targeted rewiring resulted in a significantly stronger decrease in spectral peak height than random rewiring (all $p < 0.001$, *t*-tests, Bonferroni corrected).

Supplementary Figure 2. Duplication coefficient and central peak height of empirical networks with clustering coefficient regressed out. After regressing out effects of clustering on the height of the central peak in Laplacian spectra, the

duplication coefficient remained to show a significant correlation with peak height ($r = 0.73$, $p = 0.011$).

Supplementary Figure 3. Percentage of connected most similar nodes versus central peak position. The percentage of nodes connected to their most similar peer showed to be significantly correlated with the position of the central peak in neural network spectra ($r = 0.87$, $p = 0.025$).

Table 1. Overview of the analyzed non-neural empirical networks. Detailed descriptions and references to the original sources of data are provided as Supplementary Material.

Network	Description	Number of nodes	Density (%)
Food web	Food web of Florida Bay	121	24
Words	The adjacencies of common adjectives and nouns in a novel	112	6.8
Books	Co-purchases of books on US politics	105	8.1
Dolphins	Social interactions in a community of dolphins	62	8.4
Les Miserables	Co-appearances of characters in the novel “Les Miserables”	77	8.7

Supplementary Material for the article

The role of symmetry in neural networks and their Laplacian spectra

Siemon C. de Lange, Martijn P. van den Heuvel and Marcel A. de Reus

Brain Center Rudolf Magnus, Department of Psychiatry, University Medical Center Utrecht, 3584 CX Utrecht, The Netherlands

Empirical networks

All empirical networks were transformed to undirected, unweighted networks by considering directed edges as reciprocal edges and discarding connection weights. Unless otherwise stated, all networks can be downloaded from: http://www.cise.ufl.edu/research/sparse/matrices.

Words. In this network, nodes represented the most commonly occurring adjectives and nouns in the novel "David Copperfield" by Charles Dickens and edges represented common adjacencies of these words (Newman, 2006). The words network included 112 nodes and the density was 6.8%.

Books. The books network described co-purchases of books on US politics on Amazon.com (Krebs, 2001). Nodes represented books and the edges represented frequent co-purchases of books, indicated by the "customers who bought this book also bought these other books" feature. The network included 105 books and had a density of 8.1%.

Dolphins. This social network described the relations in a community of 62 dolphins (Lusseau et al., 2003). Between frequently associated dolphins an edge was placed, resulting in a density of 8.4%.

Les Miserables. The Les Miserables network described a social network of characters from Victor Hugo's novel "Les Miserables" (Newman and Girvan, 2004). Two characters were considered to be related (and connected in the network) if the characters appeared at least once in the same chapter, giving a density of 8.7%.

Food web. The food web network described the food web of living compartments in Florida Bay during wet season [http://vlado.fmf.uni-lj.si/pub/networks/data/bio/foodweb/foodweb.htm, (Ulanowicz et al., 1998)]. The network consisted of 121 connected living compartments and had a density of 24%.

References

Krebs, V. (2001). Unpublished. Available at: http://www.orgnet.com.

Lusseau, D., Schneider, K., Boisseau, O.J., Haase, P., Slooten, E., and Dawson, S.M. (2003). The bottlenose dolphin community of Doubtful Sound features a large proportion of long-lasting

associations. Behav. Ecol. Sociobiol. 54, 396–405. doi:10.1007/s00265-003-0651-y.

Newman, M.E.J. (2006). Finding community structure in networks using the eigenvectors of matrices. Phys. Rev. E 74, 036104. doi:10.1103/PhysRevE.74.036104.

Newman, M.E.J., and Girvan, M. (2004). Finding and evaluating community structure in networks. Phys. Rev. E 69, 026113. doi:10.1103/PhysRevE.69.026113.

Ulanowicz, R.E., Bondavalli, C., and Egnotovich, M.S. (1998). Network analysis of trophic dynamics in South Florida ecosystem, FY 97: the Florida bay ecosystem. Annual Report to the United States Geological Service. Chesapeake Biological Laboratory, [UMCES] CBL 98–123.

Supplementary Material for the article

The role of symmetry in neural networks and their Laplacian spectra

Siemon C. de Lange, Martijn P. van den Heuvel, Marcel A. de Reus

Brain Center Rudolf Magnus, Department of Psychiatry, University Medical Center Utrecht, 3584 CX Utrecht, The Netherlands

In this supplementary document, we introduce further mathematical arguments supporting the empirically observed relationship between network symmetry and the central peak in Laplacian spectra of networks. First, we reproduce the proof of the proposition that two perfectly symmetric nodes in a network give rise to an eigenvalue $\lambda = 1$ in the associated normalized Laplacian spectrum (as shown, among others, by Banerjee and Jost [1]). Next, we examine network motifs exhibiting specific deviations from perfect symmetry and show these motifs to generate eigenvalues with a fixed –generally small– distance to 1. Finally, we consider motifs capturing more general forms of approximate node symmetry and show that networks possessing such a motif have a Laplacian eigenvalue close to 1.

Normalized Laplacian spectrum

Let $G = (V, E)$ be a graph described by a set V of n nodes and a set of edges E. This graph can be represented by an adjacency matrix A whose entries $A_{p,q}$ reflect the presence ($A_{p,q} = 1$) or absence ($A_{p,q} = 0$) of an edge from node p to node q. The normalized Laplacian matrix $\mathcal{L}$ of graph G has three versions:

$$\begin{aligned}\mathcal{L}_{\text{left}}(A) &= I - D^{-1}A,\\ \mathcal{L}_{\text{right}}(A) &= I - AD^{-1},\\ \mathcal{L}_{\text{symmetric}}(A) &= I - D^{-1/2}AD^{-1/2},\end{aligned}$$

where D is a diagonal matrix with $D_{p,p}$ equal to the degree d_p of node p. Since these different versions of the normalized Laplacian matrix can easily seen to be similar (in the mathematical sense), their associated eigenvectors can differ, but their eigenvalues coincide. In the following sections, we may therefore freely switch between the different definitions.

Perfect symmetry

The occurrence of perfectly symmetric nodes in a graph is related to the existence of a Laplacian eigenvalue $\lambda = 1$ by the following proposition, previously presented by Banerjee and Jost [1].

Proposition 1. *Let $G = (V, E)$ be a graph comprising two (non-adjacent) nodes p and q with identical connectivity profiles. Then the normalized Laplacian spectrum of G possesses an eigenvalue $\lambda = 1$.*

Proof. The eigenvalue equation $\mathcal{L}x = \lambda x$ for $\mathcal{L} = I - D^{-1}A$ is equivalent to $D^{-1}Ax = (1 - \lambda)x$ and hence to the following local eigenvalue equation per node:

$$\frac{1}{d_v} \sum_{u \sim v} x_u = (1 - \lambda)x_v, \text{ for all } v \in V, \tag{1}$$

where the sum is over all neighbors u of node v and d_v denotes the degree of node v. The vector x with two nonzero entries $x_p = -1$ and $x_q = 1$ can be examined as potential eigenvector. For all nodes not connected or equal to p or q, both sides of equation (1) reduce to zero. For nodes $v \in V \setminus \{p, q\}$ connected to both p and q, the equation becomes

$$\frac{1}{d_v}(x_p + x_q) = 0,$$

which is satisfied because $x_p = -x_q$. Moreover, because p and q have identical connectivity profiles, there are no nodes $v \in V \setminus \{p, q\}$ connected to only p or q. Finally, at nodes p and q, equation (1) gives

$$0 = (1 - \lambda)x_p,$$
$$0 = (1 - \lambda)x_q,$$

both of which are satisfied for $\lambda = 1$. Hence, x is an eigenvector of $\mathcal{L}$ with associated eigenvalue $\lambda = 1$. □

Approximate symmetry

In some cases, the occurrence of two non-perfectly symmetric nodes induces a Laplacian eigenvalue with a specific distance to 1. Two simple deviations from perfect symmetry are investigated:

Example 1. *Let G be a graph comprising two mutually connected nodes p and q with otherwise equal connectivity profiles. Then the Laplacian spectrum of G possesses an eigenvalue $\lambda = 1 + 1/d_p$.*

Proof. Consider the potential eigenvector x with zero entries except for $x_p = -1$ and $x_q = 1$. For nodes $v \in V \setminus \{p, q\}$ not connected to either p or q, the local eigenvalue equation (1) is zero on both sides. For nodes $v \in V \setminus \{p, q\}$ connected to both p and q, the local eigenvalue equation is

$$\frac{1}{d_v}(x_p + x_q) = (1 - \lambda)x_v,$$

which is zero on both sides because $x_p = -x_q$ and $x_v = 0$. As in the proof of Proposition 1, the symmetry assumption ensures there are no nodes $v \in$

$V \setminus \{p, q\}$ connected to only p or q. Moreover, for nodes p and q, equation (1) becomes

$$\frac{1}{d_p} x_q = (1 - \lambda) x_p,$$
$$\frac{1}{d_q} x_p = (1 - \lambda) x_q.$$

Substituting $x_p = -1$ and $x_q = 1$ gives

$$\frac{1}{d_p} = -(1 - \lambda),$$
$$-\frac{1}{d_q} = (1 - \lambda),$$

and since $d_p = d_q$, both equations are satisfied if $\lambda = 1 + 1/d_p$. □

Example 2. *Let G be a graph comprising two non-adjacent nodes p and q whose connectivity profiles are equal except that node p is connected to a leaf node r and node q is connected to a leaf node $s \neq r$. Then the normalized Laplacian spectrum of G possesses an eigenvalue $\lambda = 1 - 1/\sqrt{d_p}$ and an eigenvalue $\lambda = 1 + 1/\sqrt{d_p}$.*

Proof. Consider the potential eigenvector x with nonzero entries $x_p = 1$, $x_q = -1$, $x_r = \sqrt{d_p}$ and $x_s = -\sqrt{d_q}$. For nodes $v \in V \setminus \{p, q\}$ not connected to either p or q, both sides of the local eigenvalue equation vanish. (Note that v cannot be connected to r or s since these leaf nodes are only connected to respectively p and q. Moreover, v cannot be equal to r or s because v is not connected to p or q. That is, all involved terms x_u are equal to 0.) For nodes $v \in V \setminus \{p, q\}$ connected to both p and q (implying $v \notin \{r, s\}$), the local eigenvalue equation becomes

$$\frac{1}{d_v}(x_p + x_q) = (1 - \lambda) x_v,$$

which reduces to zero on both sides since $x_p = -x_q$ and $x_v = 0$. For node r and s, the local equations are

$$\frac{1}{d_r} = (1 - \lambda)\sqrt{d_p},$$
$$-\frac{1}{d_s} = -(1 - \lambda)\sqrt{d_q},$$

both of which are satisfied for $\lambda = 1 - 1/\sqrt{d_p}$. Also the local equations for node p and q

$$\frac{1}{d_p}\sqrt{d_p} = (1 - \lambda),$$
$$-\frac{1}{d_q}\sqrt{d_q} = -(1 - \lambda),$$

are satisfied for $\lambda = 1 - 1/\sqrt{d_p}$, so x is an eigenvector of $\mathcal{L}$ with eigenvalue $\lambda = 1 - 1/\sqrt{d_p}$. In a similar manner, one can show that vector x with nonzero entries $x_p = 1$, $x_q = -1$, $x_r = -\sqrt{d_p}$ and $x_s = \sqrt{d_q}$ is an eigenvector of $\mathcal{L}$ with eigenvalue $\lambda = 1 + 1/\sqrt{d_p}$. □

Determining explicit eigenvectors and eigenvalues for motifs showing more complex deviations from perfect node symmetry is difficult. Alternatively, the normalized Laplacian spectrum of a graph with approximately symmetric nodes can be shown to possess an eigenvalue with bounded distance to 1.

Theorem 1. *Let G be a graph with two, not mutually connected, approximately symmetric nodes p and q whose connectivity profiles differ on nodes $r_1, \ldots, r_m$ to which p but not q is connected, and nodes $s_1, \ldots, s_m$ to which q but not p is connected. Then the normalized Laplacian spectrum of G possesses an eigenvalue $1 - \epsilon \leq \lambda \leq 1 + \epsilon$, with*

$$\epsilon = \frac{1}{\sqrt{2}} \left(\frac{1}{d_p d_{r_1}} + \cdots + \frac{1}{d_p d_{r_m}} + \frac{1}{d_q d_{s_1}} + \cdots + \frac{1}{d_q d_{s_m}} \right)^{1/2}.$$

Proof. Let A and D denote the adjacency matrix and diagonal degree matrix associated with graph G and define the matrix $M := D^{-1/2} A D^{-1/2}$. Then $M^2 = D^{-1/2} A D^{-1} A D^{-1/2}$ is a real symmetric matrix, hence, according to the min-max theorem, its smallest eigenvalue ν_1 is given by:

$$\nu_1 = \min_{\|x\|=1} x^\intercal M^2 x. \tag{2}$$

Since M is diagonalizable and $M^2 = (I - \mathcal{L})^2$, one can easily check that there must exist an eigenvalue λ of $\mathcal{L}$ such that $\nu_1 = (1 - \lambda)^2$. According to equation (2), this eigenvalue λ satisfies:

$$(1 - \lambda)^2 = \min_{\|x\|=1} x^\intercal M^2 x. \tag{3}$$

If nodes p and q would be perfectly symmetric, matrix M^2 would possess an eigenvalue $\nu = 0$ associated with the (normalized) eigenvector x with nonzero entries $x_p = 1/\sqrt{2}$ and $x_q = -1/\sqrt{2}$. In the current case in which nodes p and q are only approximately symmetric, vector x can be used as a best guess to bound (3), giving

$$\begin{aligned} (1 - \lambda)^2 &\leq x^\intercal M^2 x \\ &= \frac{1}{2} \left(M^2_{p,p} + M^2_{q,q} - M^2_{p,q} - M^2_{q,p} \right). \end{aligned}$$

The expression $M^2_{p,p} + M^2_{q,q} - M^2_{p,q} - M^2_{q,p}$ can be expanded by writing

$$\sum_{v \in V} M_{p,v} M_{v,p} + \sum_{v \in V} M_{q,v} M_{v,q} - \sum_{v \in V} M_{p,v} M_{v,q} - \sum_{v \in V} M_{q,v} M_{v,p}.$$

Rearranging terms and using the definition of M gives

$$\sum_{v \in V} (M_{p,v}M_{v,p} + M_{q,v}M_{v,q} - M_{p,v}M_{v,q} - M_{q,v}M_{v,p}) \tag{4}$$

$$= \sum_{v \in V} \frac{A_{p,v}}{d_p d_v} + \frac{A_{q,v}}{d_q d_v} - \frac{A_{p,v}A_{q,v}}{d_v\sqrt{d_p d_q}} - \frac{A_{p,v}A_{q,v}}{d_v\sqrt{d_p d_q}}. \tag{5}$$

For simplicity, we assumed that nodes p and q have equal degree, hence this expression becomes

$$\sum_{v \in V} \frac{A_{p,v}}{d_p d_v} + \frac{A_{q,v}}{d_p d_v} - 2\frac{A_{p,v}A_{q,v}}{d_v d_p}.$$

Nodes v that are not connected to either p or q, or to both p and q make no contribution to the summation. For nodes $v \in \{r_1, \ldots, r_m\}$, q and v are not connected and the expression reduces to $\frac{1}{d_p d_v}$. Similarly, the contribution of nodes $v \in \{s_1, \ldots, s_m\}$ equals $\frac{1}{d_q d_v}$. Combining these observations, we find that there exists an eigenvalue λ of $\mathcal{L}$ such that

$$(1-\lambda)^2 \leq \frac{1}{2}\left(\frac{1}{d_p d_{r_1}} + \cdots + \frac{1}{d_p d_{r_m}} + \frac{1}{d_q d_{s_1}} + \cdots + \frac{1}{d_q d_{s_m}}\right) \square$$

Discussion

The investigated examples of approximate node symmetry and the presented theorem for more general cases of non-perfect node symmetry provide mathematical arguments supporting the existence of a relationship between approximate node symmetry and the occurrence of Laplacian eigenvalues with a fixed or bounded distance to 1, shaping the central spectral peak. We should, however, stress that the here presented observations are not intended to provide a complete theoretical foundation for our empirical findings. First of all, Theorem 1 considers only a single pair of approximately symmetric nodes, which are additionally assumed to have equal degree. Because the expression in equation (5) cannot easily be simplified further if nodes p and q differ in degree, the bound for eigenvalues associated with approximately symmetric nodes with different degrees is likely to be of a more complex form. Moreover, additional arguments will be required to show how the occurrence of multiple couples of approximately symmetric nodes results in multiple eigenvalues with a bounded distance to 1.

Second, it is not clear how well the bound presented in Theorem 1 approximates the actual eigenvalue. In this context, it is worth to note that for the two investigated example deviations from perfect symmetry, the eigenvalue bound coincides with the expression for the exact eigenvalues. In the first example of two symmetric but mutually connected nodes p and q, the eigenvalue bound is

$$\epsilon = \frac{1}{\sqrt{2}}\left(\frac{2}{d_p d_q}\right)^{1/2} = \frac{1}{d_p},$$

which coincides with the exact eigenvalue $\lambda = 1 + 1/d_p$ associated with this motif. The second example describes two non-adjacent nodes p and q with connectivity profiles that are equal except that node p is connected to a leaf node r and node q is connected to a leaf node s. The bound for this example is

$$\epsilon = \frac{1}{\sqrt{2}} \left(\frac{2}{d_p} \right)^{1/2} = \frac{1}{\sqrt{d_p}},$$

precisely giving the distance between 1 and the exact eigenvalues $\lambda = 1 - 1/\sqrt{d_p}$ and $\lambda = 1 + 1/\sqrt{d_p}$ determined for this motif.

Third, the here presented results express eigenvalues in terms of the degree of the nodes involved, whereas our empirical findings presented in the main text use the matching index-based duplication coefficient as a measure of symmetry. These two concepts are, however, related. Assuming for simplicity that all nodes have degree d (i.e., G is regular), the expression for ϵ reduces to

$$\epsilon = \frac{\sqrt{h_{p,q}}}{d\sqrt{2}},$$

where $h_{p,q}$ is the Hamming distance between nodes p and q, representing the number of nodes to which their connectivity differs. Similarly, also the matching index between nodes p and q is determined by degree and Hamming distance, albeit in a different manner:

$$\mathrm{MI}(p,q) = \frac{d - h_{p,q}/2}{d + h_{p,q}/2} = \frac{1 - \epsilon^2 d}{1 + \epsilon^2 d}.$$

Hence, the existence of node pairs with a small bound ϵ results in a high duplication coefficient (with the above expression being maximum for $\epsilon = 0$) and Laplacian eigenvalues close to 1.

References

[1] A. Banerjee and J. Jost, "On the spectrum of the normalized graph Laplacian," *Linear Algebra and its Applications*, vol. 428, pp. 3015–3022, jun 2008.

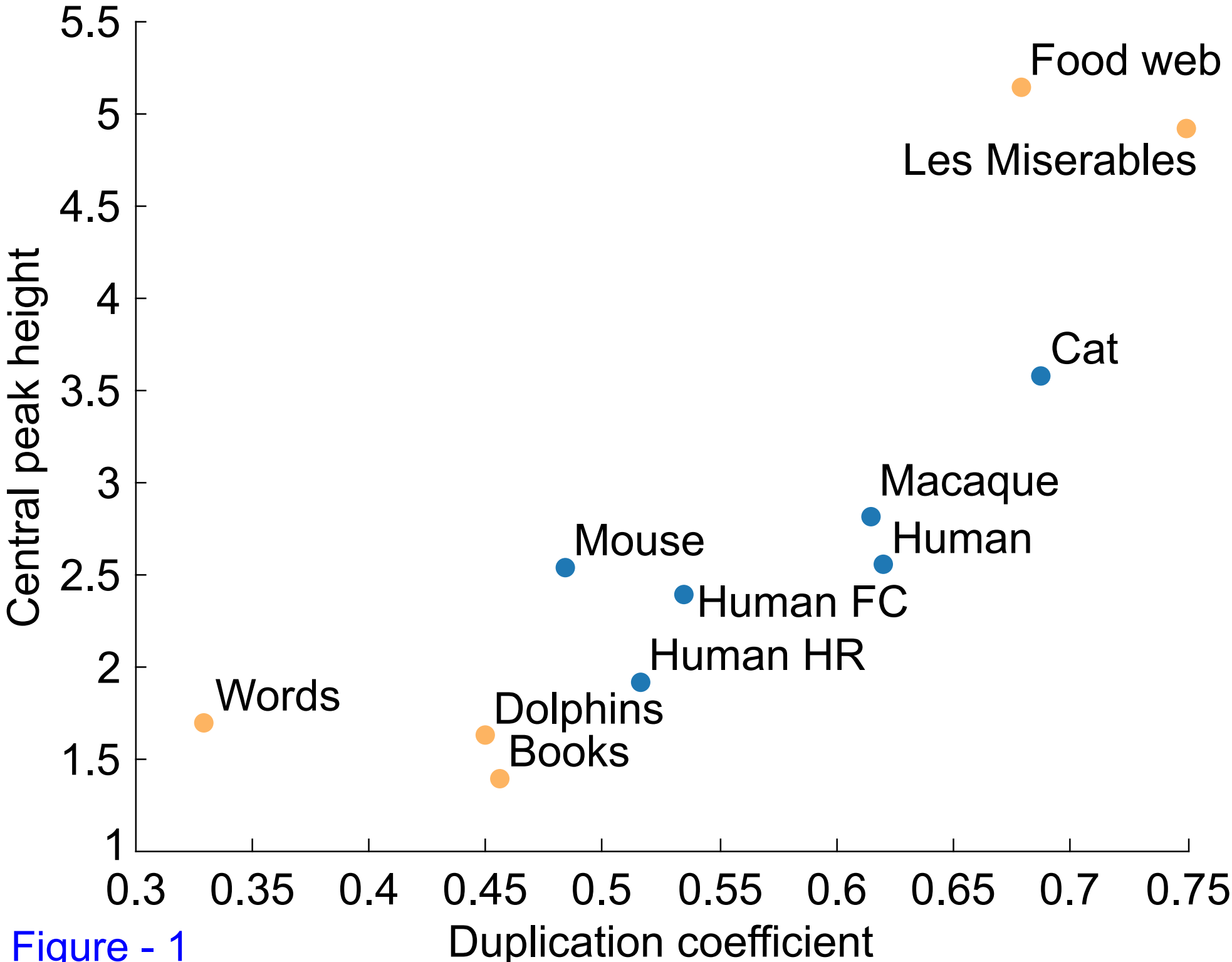
Food web
Les Miserables
Cat
Macaque
Human
Mouse
Human FC
Human HR
Words
Dolphins
Books
Central peak height
Duplication coefficient
5.5
5
4.5
4
3.5
3
2.5
2
1.5
1
0.3
0.35
0.4
0.45
0.5
0.55
0.6
0.65
0.7
0.75

Figure - 1

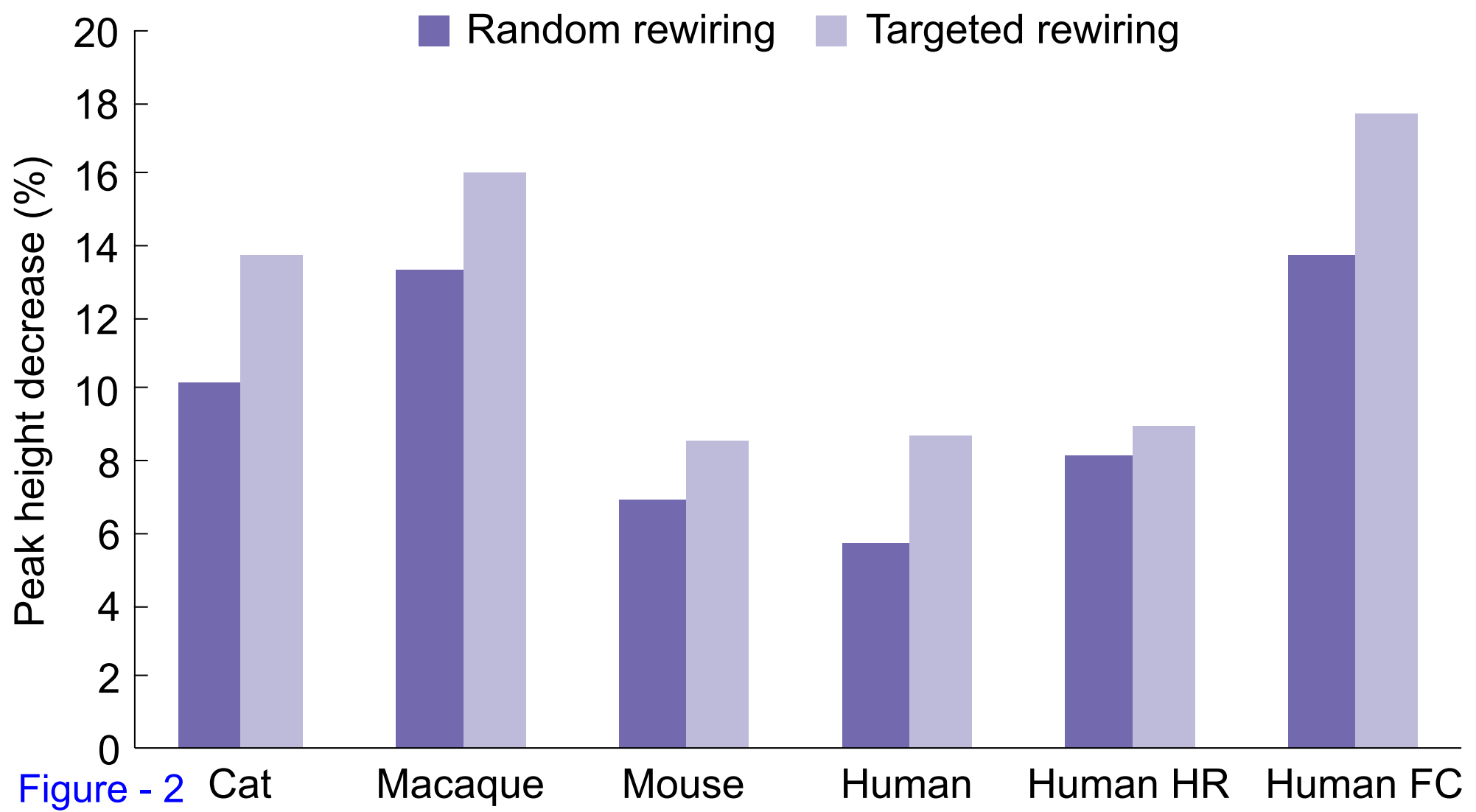
Random rewiring
Targeted rewiring
Peak height decrease (%)
0
2
4
6
8
10
12
14
16
18
20
Cat
Macaque
Mouse
Human
Human HR
Human FC

Figure - 2

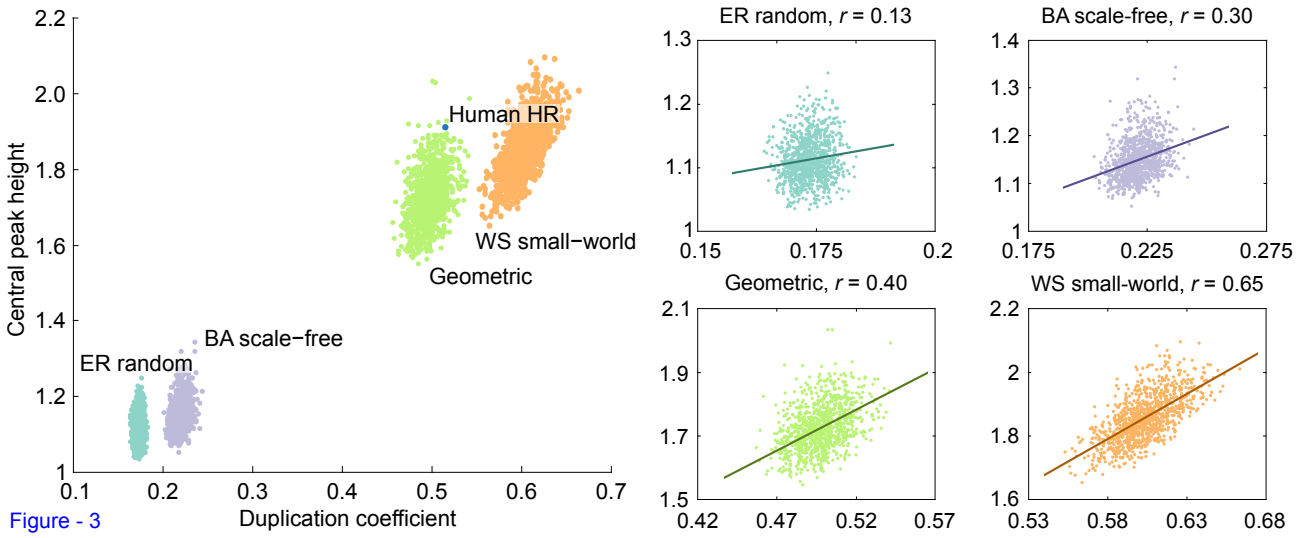
Central peak height
Duplication coefficient
Human HR
WS small−world
Geometric
BA scale−free
ER random
ER random, r = 0.13
BA scale-free, r = 0.30
Geometric, r = 0.40
WS small-world, r = 0.65

Figure - 3

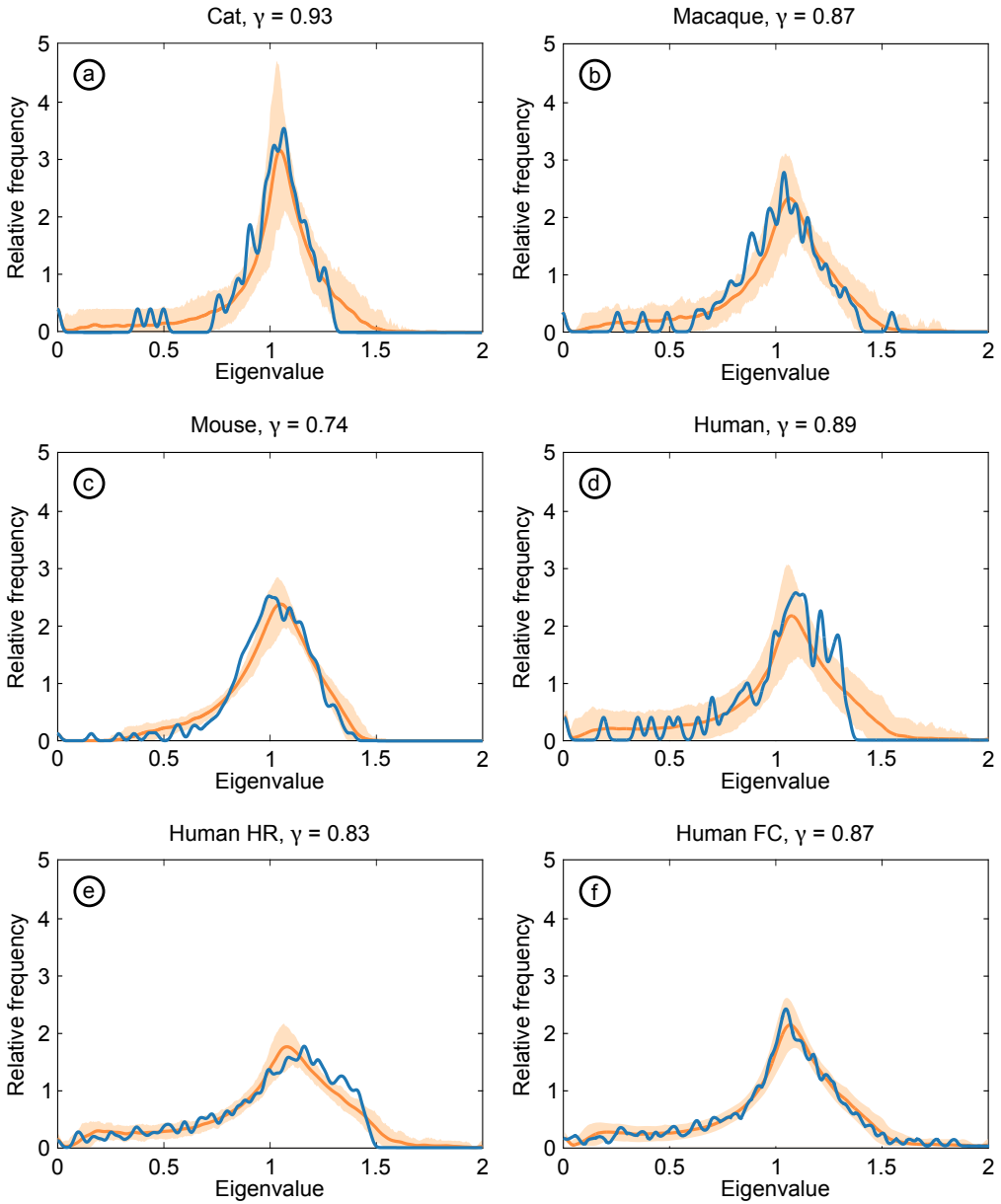
Cat, γ = 0.93
Macaque, γ = 0.87
Mouse, γ = 0.74
Human, γ = 0.89
Human HR, γ = 0.83
Human FC, γ = 0.87
a
b
c
d
e
f
Relative frequency
Eigenvalue
0
0.5
1
1.5
2
0
1
2
3
4
5
Neural network
Duplication model
95% Confidence interval

Figure 4

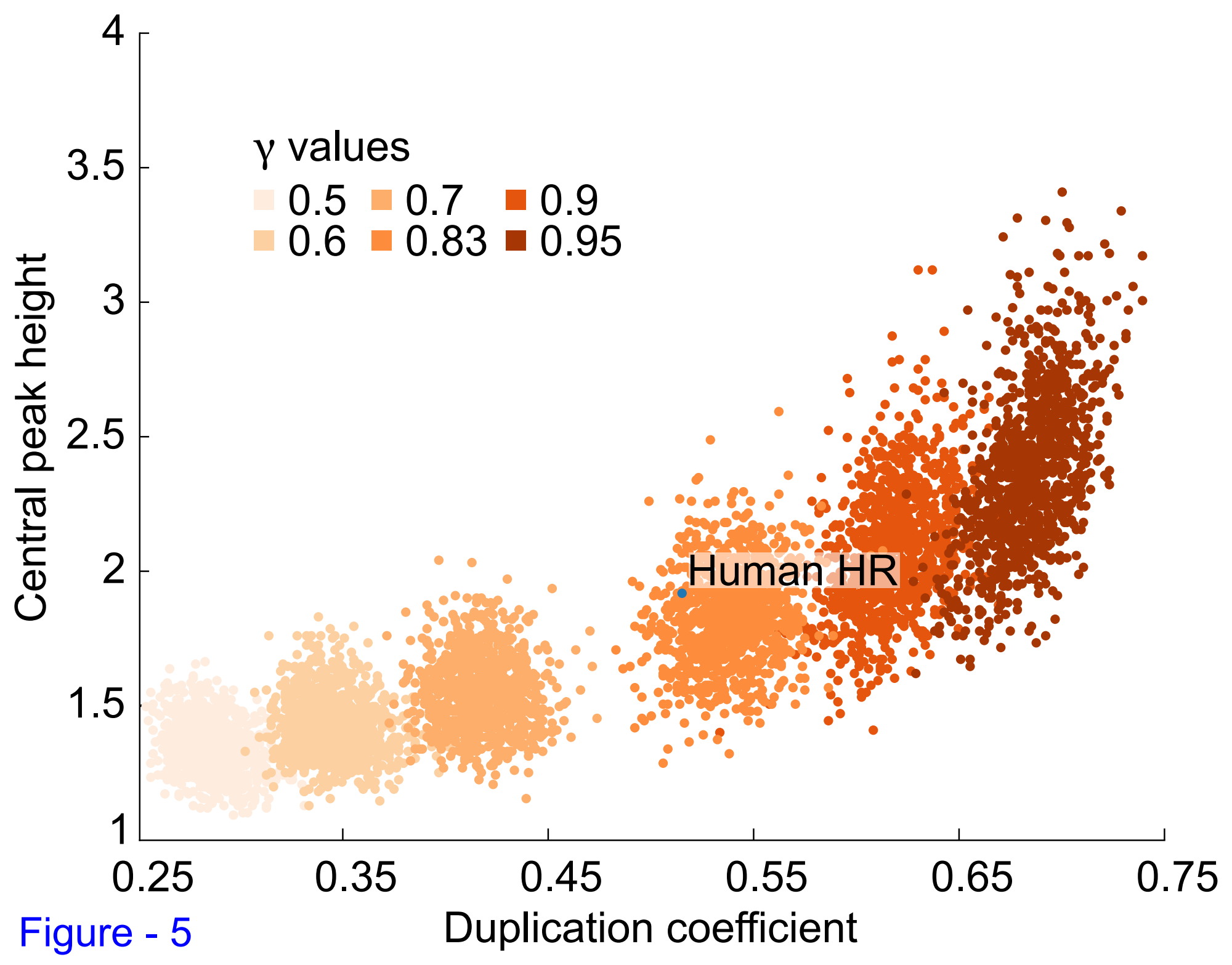
γ values
0.5
0.6
0.7
0.83
0.9
0.95
Human HR
Central peak height
Duplication coefficient
4
3.5
3
2.5
2
1.5
1
0.25
0.35
0.45
0.55
0.65
0.75

Figure - 5

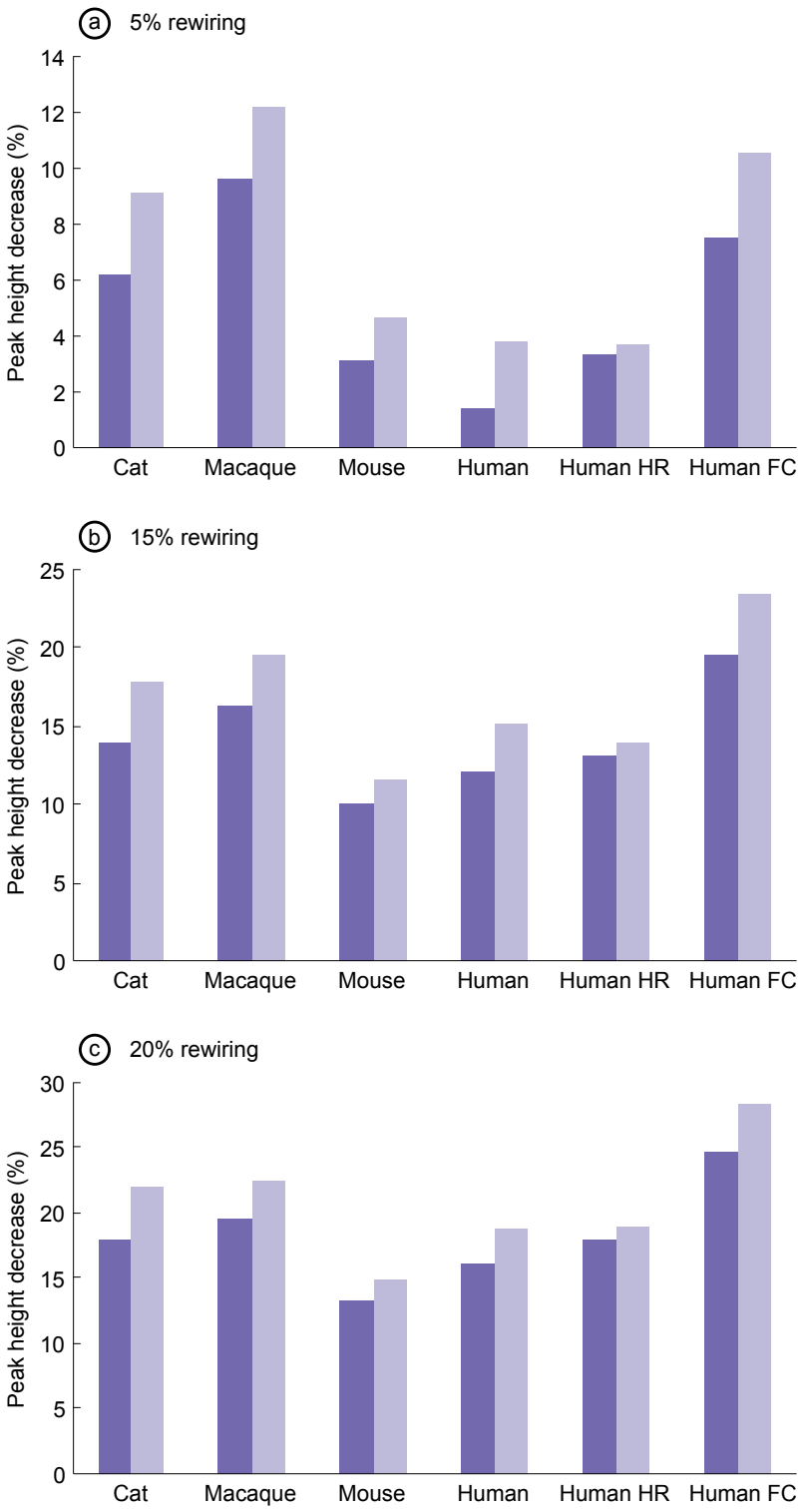
a
5% rewiring
Peak height decrease (%)
0
2
4
6
8
10
12
14
Cat
Macaque
Mouse
Human
Human HR
Human FC
b
15% rewiring
Peak height decrease (%)
0
5
10
15
20
25
Cat
Macaque
Mouse
Human
Human HR
Human FC
c
20% rewiring
Peak height decrease (%)
0
5
10
15
20
25
30
Cat
Macaque
Mouse
Human
Human HR
Human FC
Random rewiring
Targeted rewiring

Figure SI - 1

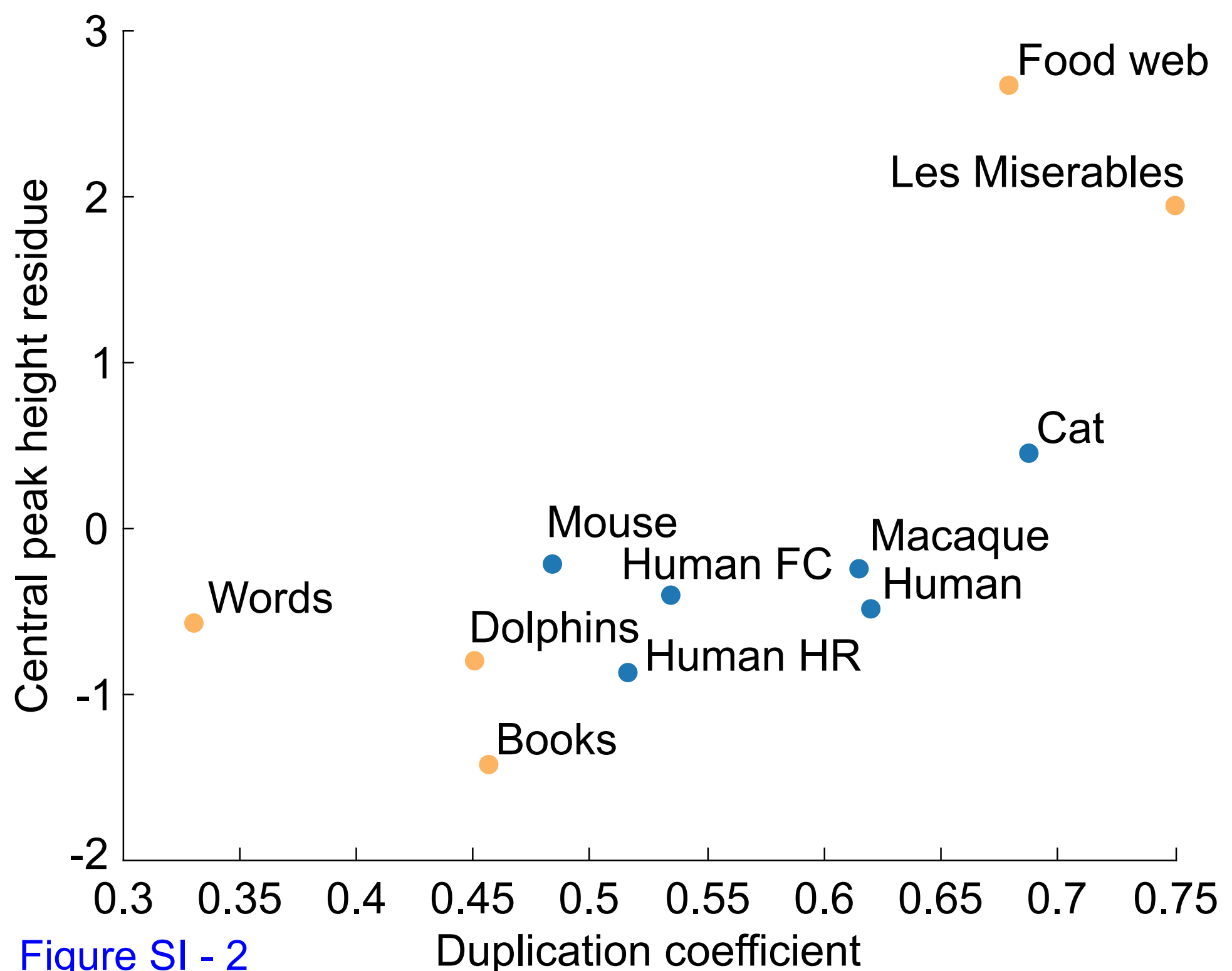
Food web
Les Miserables
Cat
Mouse
Macaque
Human FC
Human
Words
Dolphins
Human HR
Books
Central peak height residue
Duplication coefficient
3
2
1
0
-1
-2
0.3
0.35
0.4
0.45
0.5
0.55
0.6
0.65
0.7
0.75

Figure SI - 2

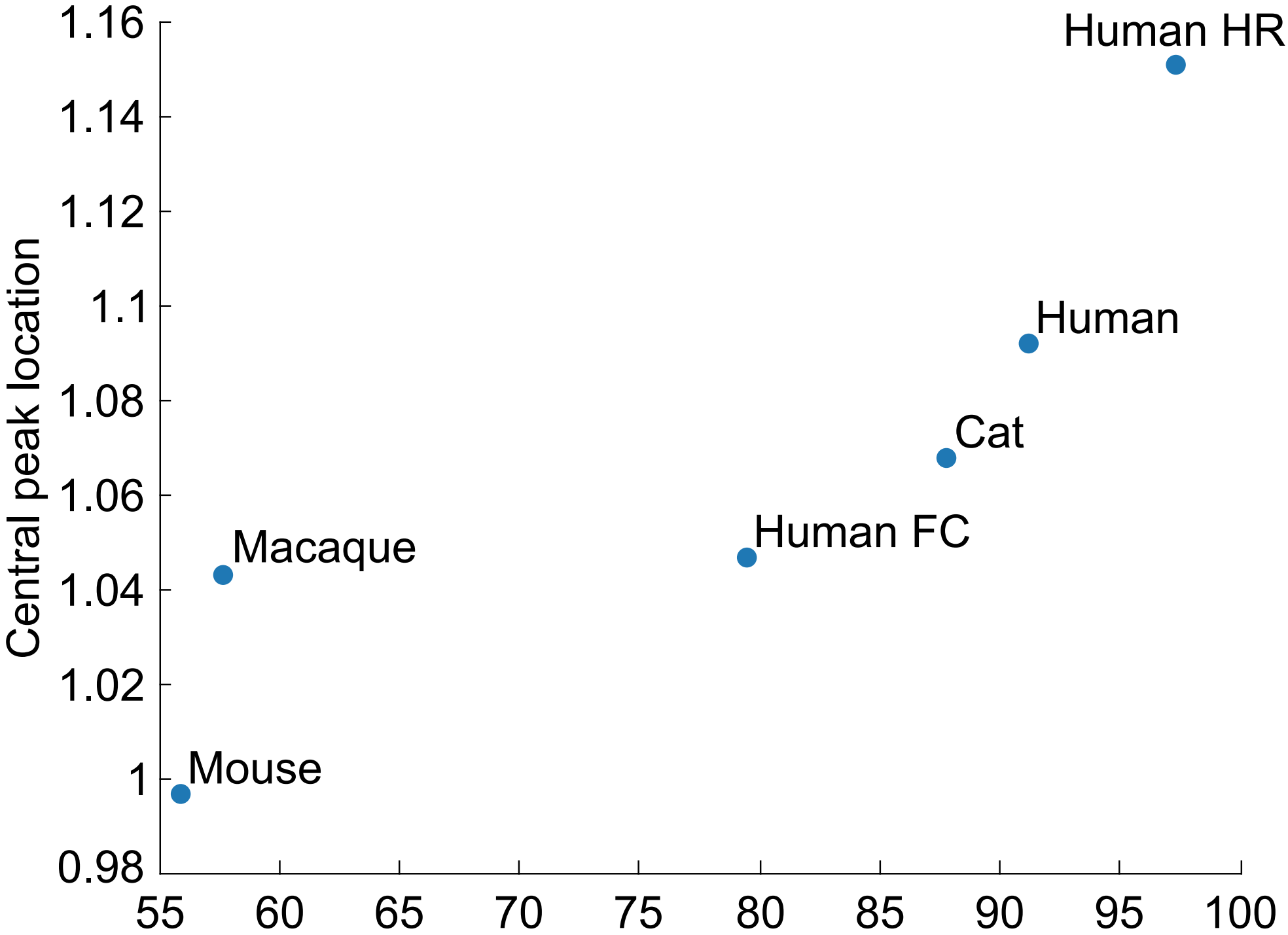
Human HR
Human
Cat
Human FC
Macaque
Mouse
Central peak location
1.16
1.14
1.12
1.1
1.08
1.06
1.04
1.02
1
0.98
55
60
65
70
75
80
85
90
95
100
Percentage of connected most similar nodes

Figure SI - 3